University of Wisconsin-Madison

# Institute for Research on Poverty

## Discussion Papers

Richard Campbell
Elizabeth Mutran

ANALYZING PANEL DATA IN STUDIES OF AGING: APPLICATIONS OF THE LISREL MODEL

DP #650-81

# Analyzing Panel Data in Studies of Aging:

# Applications of the LISREL Model

Richard T. Campbell
Duke University
and
Institute for Research on Poverty
University of Wisconsin, Madison

Elizabeth Mutran
University of Akron

April 1981

Preparation of this paper was facilitated by a Career Development Award to Campbell from the Center for the Study of Aging and Human Development, Duke University; by NIMH Grant 32348; and by funds granted to the Institute for Research on Poverty at the University of Wisconsin, Madison, by the Department of Health and Human Services pursuant to the provisions of the Economic Opportunity Act of 1964. We are indebted to Robert N. Parker, Peter Mossel, and Gerald Marwell for helpful comments on an earlier draft.

## ABSTRACT

The LISREL model, recently developed by Joreskog and his colleagues, is of great importance to students of aging and human development. Essentially, LISREL unites factor analysis and structural equation modeling. It is extremely general, permitting a wide variety of nonrecursive structural equation models while also permitting complex models involving measurement error. Thus, LISREL speaks to the psychometric concerns of those who measure the same variable over time and to the modeling concerns of those who wish to allow for lagged effects, feedback loops, and the like. This paper describes the LISREL model in nontechnical terms, points the reader to the more technical literature, and provides an extended example of a three-wave, two-variable model, with and without additional exogenous and endogenous variables. The example shows how causal effects can be estimated and how these estimates are affected by assumptions regarding measurement error.

# Analyzing Panel Data in Studies of Aging: Applications of the LISREL Model

## INTRODUCTION

Research on aging frequently involves repeated measures on the same sample of individuals. The panel design, in which measures are obtained at fixed intervals, with retrospective measures of intervening events, has become commonplace in gerontology. Perhaps the most familiar example for gerontologists is the Longitudinal Retirement History Survey (LRHS), now being conducted by the Social Security Administration (Irelan, 1972). It consists of 5 waves of measurement taken biennially between 1969 and 1978, from more than 10,000 persons. There are several other similar national panel surveys, including the National Longitudinal Surveys of Labor Force Participation (Parnes, 1975), the Panel Study of Income Dynamics (Morgan et al., 1972) and numerous other more local studies. All of the surveys just mentioned, and many others of a similar nature, are in the public domain and easily accessible to researchers in gerontology.

Despite the plenitude of data, it is fair to say that the number of published analyses which make use of more than two waves of data is extremely small. Most researchers, even those with a fair amount of statistical sophistication, have found the analysis of multiwave panel data to be a formidable challenge, and rightly so. Until recently, easily accessible techniques for handling multiwave data have not been available. Researchers trained in psychology, as many gerontologists are, have had the advantage of the conceptual power to be found in the analysis of variance paradigm. Multiwave data can be seen as a problem in repeated measures, and recent developments in the application of multivariate analysis of variance (MANOVA) to repeated measures (Bock, 1979), particularly growth curve approaches, have been helpful. Using this approach one could, for example, look at changes in measured intelligence over time, comparing curves across such groups as those defined by sex and cohort. Although MANOVA is most commonly applied to physiological or cognitive variables, there is no statistical reason why it cannot be applied to, say, changes in income levels over time or to other similar measures.

However, as Baltes and Nesselroade (1979) note, "non-normative" events--i.e., those which do not follow a biological or social time clock--become increasingly important in the determination of life-course phenomena late in the life span. Thus, education follows a generally predictable time course, and income levels early in the career generally follow carefully defined schedules corresponding to increasing education and experience. Later in life, however, social variables such as marital status and income level or psychological variables such as life satisfaction respond to a host of non-normative events--illnesses, marital disruption, migration, and other unpredictable influences. Analytically, this requires us to pay attention to the intervening variables which might cause fluctuation in the time path of a variable. For example, yearly panel data on life satisfaction are not particularly informative unless we know what happened between the waves of measurement; did the subject become a widow or widower, did that person leave the labor force, or enter the empty-nest phase, or undergo other change? If so, the analysis should take this into account. Note that the events occur at varying points in time, and occur for some individuals but not for others.

The MANOVA approach is not particularly good at handling this kind of problem, but intervening variables are rather easily handled within another analytic tradition--path analysis. The remainder of this paper provides a nontechnical introduction to advanced causal modeling techniques for panel data, based on path analytic and factor analytic concepts. Readers who are at least broadly familiar with these concepts should be able to read the exposition with little difficulty; others may wish to read introductory material first. A good place to start, particularly for those with a background in psychology, is Kenny's Correlation and Causality (1979). The SPSS Manual (Nie et al., 1975) also contains adequate descriptive information.[1]

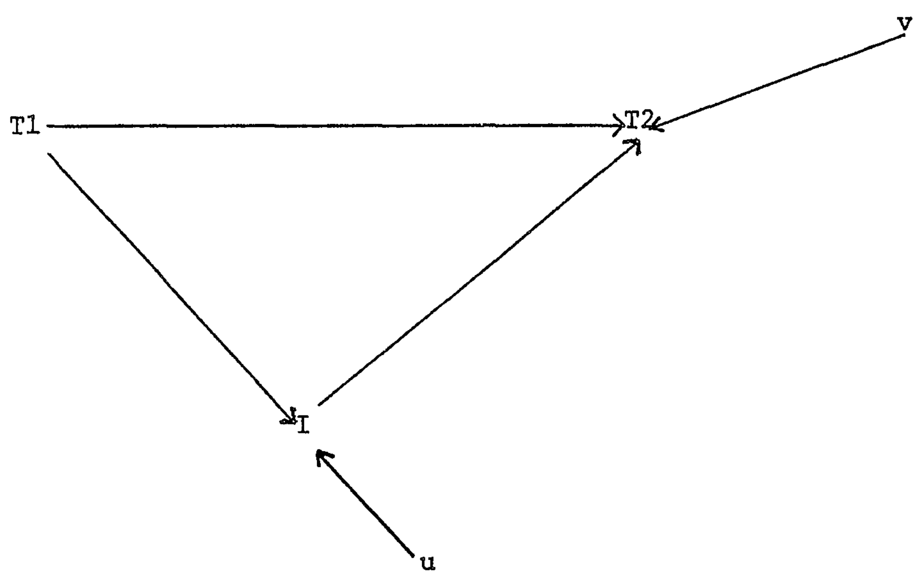
## PATH ANALYSIS OF PANEL DATA

### Two-Wave Models

Researchers trained in path analysis have little difficulty with the problem of intervening variables, but, as we shall see, they face other equally serious problems. In the simplest case of two waves of measurement, shown in Figure 1, the path analytic approach is clearly effective. In this generic diagram, T1 and T2 are repeated measures of an unspecified attitudinal variable. The letter I indicates an intervening variable, assumed to occur between the waves of measurement and ascertained on a recall basis during the T2 interview. The diagram follows the usual conventions of path analysis; the single-headed unidirectional arrows specify causal effects. The path from T1 to T2 specifies that we expect T1 to have a causal impact on T2, and the arrow from I to T2 implies that I has an impact on T2 as well. The variables u and v are residual variables introduced to account for all variance in I and T2 not accounted for by their measured determinants. This approach, in agreement with the

FIGURE 1

A SIMPLE TWO-WAVE MODEL
WITH AN INTERVENING VARIABLE



perspective advanced by Cronbach and Furby (1972), avoids the direct calculation of a change score, instead concentrating on the regression or path coefficient linking T2 with T1. This coefficient is estimated net of (controlling for) the intervening variable. The path analysis approach allows us to estimate the stability of the dependent variable over time, the extent to which change in the dependent variable is a function of the intervening variable, and the degree to which changes in T1 are transmitted to the dependent variable via the intervening variable.[2]

This basic approach to two-wave models has been used in a variety of applications, with extensions to multiple intervening variables, multiple controls, and more causal structures involving the (retrospectively measured) intervening variables. The path analytic approach, while extremely useful and informative, is not without problems, however, even in the simple two-wave case. In using the technique one is making two important assumptions in addition to those made in any regression analysis (Hanushek and Jackson, 1977, pp. 45-59); both assumptions are almost certainly violated. The first assumption is that the variables are measured perfectly. Errors in exogenous variables (i.e., those with reliabilities of less than one) will bias coefficients in path models. Thus, if both T1 and T2 are subject to random error, the estimate of stability is almost certainly biased downward, as is the estimate of $R^2$. The second assumption is that errors in variables are not correlated over time. For example, if respondents tend systematically to underestimate T1 and do the same on T2, the errors of measurement at T1 and T2 will be correlated, and the estimated stability will be biased.[3]

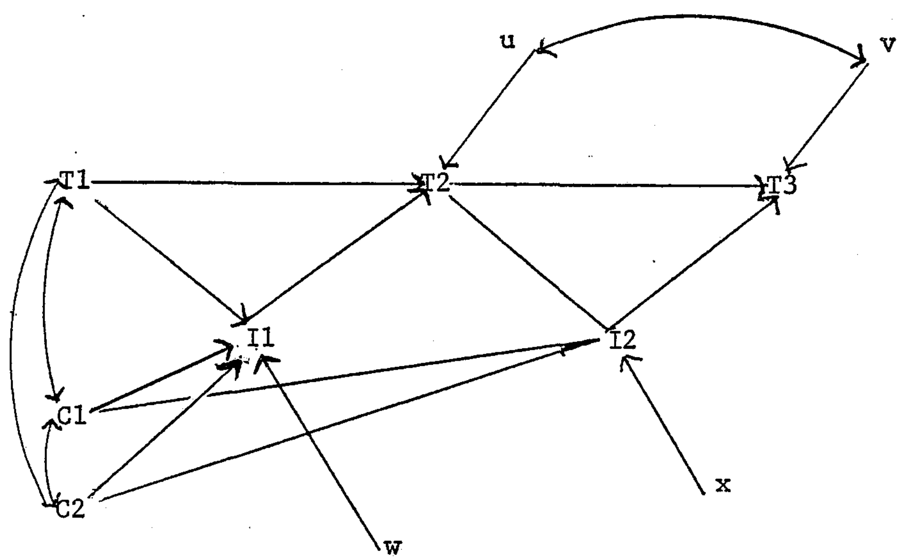
## Three-Wave Models

Disregarding these problems for a moment, let us look at the path analysis of a three-wave model. Figure 2 shows a generic three-wave, two-variable model. The notation follows the conventions of Figure 1. T1, T2, and T3 represent a specific variable measured at three points in time, I1 and I2 represent another variable that is taken to intervene between T1 and T2 and between T2 and T3 respectively. C1 and C2 are control variables. Note that we assume that the effect of T1 on T3 is indirect, via T2, I1, and I2--that is, there is no path from T1 to T3. Similarly, we assume that C1 and C2 affect T2 and T3 only through their correlation with T1 or via I1 and I2. Finally, note the curved arrow connecting the residuals at T2 and T3, indicating that we expect correlated errors in equations. Correlated error will come about if (a) the same causal variables are omitted from both equations, or (b) measurement phenomena, such as similar biases in question wording, occur at both time points. Both of these problems are likely to occur in longitudinal data.

In the case of simple path analysis one obtains a solution to recursive models by solving a set of regression equations. In Figure 1, for example, the solution was obtained by first regressing I on T1 and then regressing T2 on I and T1. In the more complex case, it can be shown that, in the presence of correlated errors in equations, simple regression analysis will result in biased and inconsistent coefficients, that is, the sample estimates will not converge to the population parameter as N increases. Moreover, if the omitted paths (such as that from T1 to T3) are in fact not zero, the estimates will be biased even if there is not correlated error. On the other hand, if we try to estimate the path from

FIGURE 2

A MORE COMPLEX
THREE-WAVE MODEL



T1 to T3 we will encounter other statistical problems because T1, T2, and T3 are, in all likelihood, highly correlated.

What we need is a statistical approach that lets us take correlated error explicitly into account while retaining the capability of setting particular coefficients to zero (or perhaps requiring that the estimates of two coefficients be equal). There are statistical solutions available for these problems (and for others engendered by longitudinal data), but they have not been of a general nature; each has required a separate, and typically complex, algebraic attack. Within the past few years, however, there has emerged an approach to structural equation models which has many desirable characteristics when applied to longitudinal data, which is quite general, and which also is comparatively simple to use. Developed by Karl Jöreskog and his students, (Jöreskog, 1969, 1970, 1979; Jöreskog and Sorbom, 1978) the basic approach is referred to as LISREL (Linear Structural Relationships). LISREL refers both to a particular statistical technique and to the computer program used to generate estimates.

All basic statistical references along with references to numerous applications of LISREL can be found in the user's manual (Version IV, Jöreskog and Sorbom, 1978). An excellent introduction to the use of LISREL for longitudinal data analysis, from which the present paper draws heavily, is Wheaton et al. (1977). A broader discussion of LISREL in the context of other approachs to the analysis of longitudinal data can be found in Schaie and Herzog (in press).

In the remainder of this paper we discuss the approach in some detail, attempting to show how it is applicable to the model shown in Figure 2 and to a wide variety of other models. LISREL deals with models of this kind at the same time that it permits one to deal with problems of measurement error in a statistically appropriate way. The LISREL program produces estimates of the various coefficients in the model, computes their standard errors, and calculates a goodness-of-fit statistic between the variance-covariance matrix implied by the model and that observed in the sample.

Within a period of ten years LISREL has emerged as a standard analytic technique for structural equation models. Unfortunately, much of the literature which presents the model is quite technical. The purpose of this paper is to show the power of the LISREL approach for the analysis of longitudinal data while avoiding much of the technical detail. We do not mean to imply that this detail is unnecessary for applications of the model, it is indeed necessary. LISREL is perhaps more subject to misapplication than many other statistical techniques, and the reader is cautioned that this paper cannot serve as the sole basis for its successful use. At the same time, we have found that many researchers are overwhelmed by LISREL's generality and its rather difficult notation. Our basic goal in this paper is to provide a conceptual overview of what LISREL does and how it does it. With this material in hand, we hope that the reader will find the more advanced presentations somewhat easier to follow.

In order to demonstrate the advantages of the LISREL perspective and to develop it within a framework that is intelligible to gerontologists, we present an analysis of the determinants of income satisfaction and changes therein using three waves of data from the LRHS. We begin with a very simple model, one which looks only at the stability of income satisfaction (measured with two indicators) over time. Although we restrict ourselves to three waves of data in this example, there is no reason why we could not extend the model to four, five, or more measurement points.

After discussing the simple model of stability, we turn to a more complex model which looks at the determinants of change in income satisfaction, using information from the LRHS on health status (measured with three indicators); we also use self-reported income at wave 1 (one indicator), and the number of times the respondent was hospitalized in 1970 (one indicator, referring to the time intervening between waves one and two). Thus, we have two variables, income satisfaction and health, in parallel over time, a base line variable (actual income), and an intervening variable (times hospitalized). We do not intend the analysis to be a full and complete substantive analysis of changes in health and income satisfaction--other variables would have to be included and many more models would have to be explored--but the analysis is sufficiently close to actual practice so that the full flavor of the technique can be appreciated.

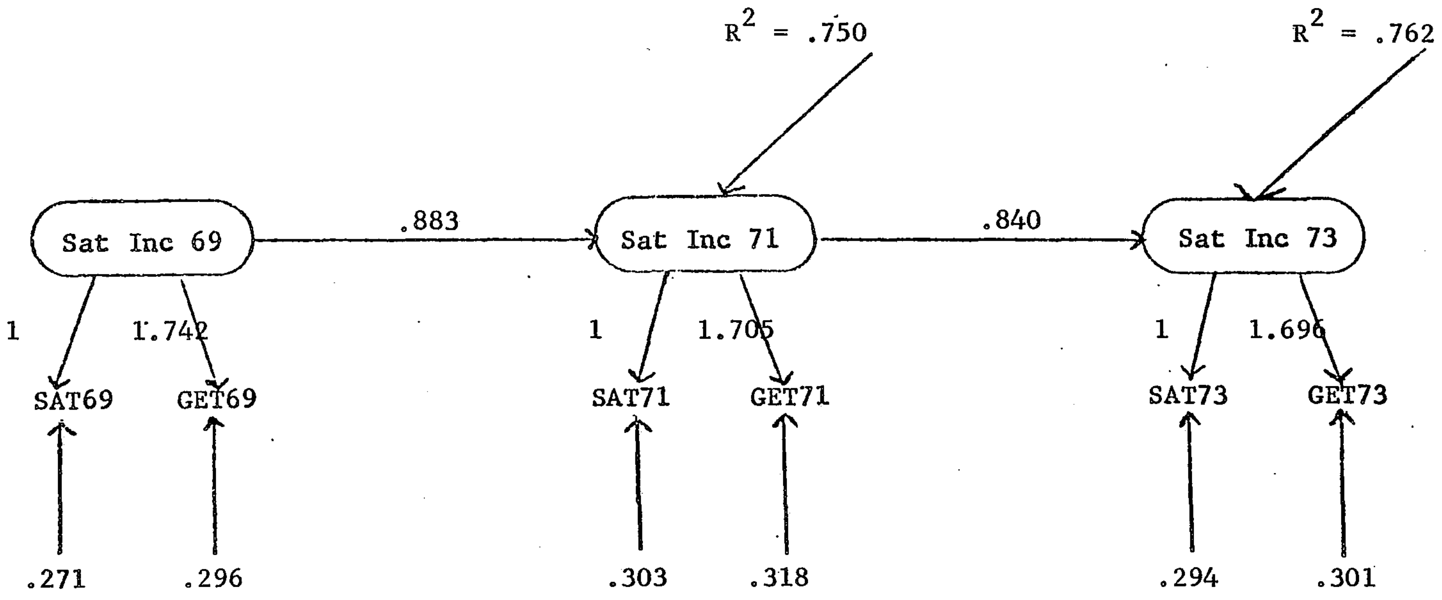
## A THREE-WAVE MODEL FOR INCOME SATISFACTION

### Causal Models and Confirmatory Factor Analysis

Figure 3 presents a path diagram relating the three measures of income satisfaction over time. Although this model is of no great substantive interest, it has the advantage of being simple while demonstrating a number of LISREL characteristics. Note first that the income satisfaction measure, represented by the boxes, is taken to be an unmeasured construct with observable indicators. The two indicators are "satisfaction with the way one is living" (SAT) and "ability to get along on income" (GET). The model assumes that these indicators are related to the underlying construct, that is, that the variances and covariances of

FIGURE 3

A SIMPLE THREE-WAVE REPEATED MEASUREMENT MODEL




Note: All coefficients are unstandardized. See Appendix A and Table A.1 for definitions and variances of the variables.
$\chi^2 = 491.8$; df = 7.

the indicators are due to their common (unmeasured) cause. In other words, we are assuming that the indicators have a particular factor structure. In fact, LISREL is a special case of confirmatory factor analysis (CFA). In CFA, one posits a specific factor structure and tests to see if that structure is congruent with the observed data. LISREL extends this one step further to allow for a causal structure on the factors.[4] The CFA model was developed by Jöreskog (1969) and was rapidly extended to the more complete causal formulation just described (1970, 1979).

Since this concept is crucial to understand what follows, it bears repeating in another way. We can think of the six measures of income satisfaction as indicators of three underlying factors. The three factors are simply the unmeasured construct "income satisfaction" at each of the three time points. We assume, of course, that these three constructs are nonindependent--i.e., that there is some stability in the ordering of individuals on the dimension of income satisfaction. If we were doing conventional factor analysis we would want an oblique (correlated) structure. An oblique solution would produce a matrix of factor intercorrelations, and that matrix, like any correlation matrix, could be subjected to causal analysis. LISREL performs these two tasks simultaneously; that is, it determines the appropriate estimates of relationships between observables and factors and then determines the estimates of causal relationships among the factors, given the user's specification of a model.[5]

Measurement models and causal models. In thinking about this dual aspect of the LISREL model, it will be useful to distinguish between the measurement part of the model and the causal or structural equation part.

The measurement model is indicated by the arrows leading from boxes in Figure 3 to the observable variables; the causal aspects are indicated by the arrows linking the boxes. Actually, the measurement part of the model also involves a causal hypothesis common to all factor analytic and classical test-score models. The arrow leading from "factor" to observable indicates the assumption that the variances and covariances among the observed variables are due to their common sources in unmeasured constructs. The coefficients linking the observables to constructs are therefore regression coefficients of observables on true scores.[6] The relationships between observables and factors or constructs is not perfect, however, and the arrow leading to each observable from below indicates variance in observables not accounted for by the factor. In classical factor analytic terms, this is equivalent to specific or unique variance, not accounted for by the factor(s), plus variance due to measurement error.

Covariance vs. correlation matrix. The numbers in Figure 3 will surely give many readers pause, since paths leading from factors to observables are greater than 1. The reason is that these are unstandardized coefficients obtained from a covariance matrix solution. Analyzing the covariance matrix allows one to depart from the artificial constraint that all variables have the same variance. In panel data, it is not uncommon to see the variance of variables increasing or decreasing over time in predictable ways. To ignore this phenomenon throws away useful information. Consequently, most LISREL analyses, whether of panel data or not, analyze the covariance matrix, though it is certainly possible to analyze the correlation matrix. Although correlational analyses would be more familiar to many readers, it would then be more difficult to turn to the primary literature, which emphasizes covariance solutions; hence we will use the unstandardized approach. (For a discussion of those issues see Heise, 1969; Wiley and Wiley, 1970.)

In conventional factor analyses, one typically assumes that the factors are standardized to a mean of zero and variance of one. This same convention could be followed in LISREL even if one were analyzing a covariance rather than a correlation matrix. However, since the metric of the factors is arbitrary, the factor loadings are only estimable up to a constant of proportionality. We can fix the variance of the factors, or we can deal with this problem by fixing one of the loadings to a specific value, thus forcing the program to estimate all other loadings on that factor relative to the fixed variable. For example, in Figure 3 the path coefficient linking SAT69 to its factor is fixed to 1.0 and the other variable on that factor (GET69) is free to vary relative to the fixed coefficient. The estimated coefficient is 1.742. Alternatively, we might have fixed the variance of the factor to a specific value, say 1. Had that been done in this case, the factor loadings would have been .448 and .780, reflecting the same relative values, since .780/.448 = 1.741.

The measurement model. With these preliminaries in hand, we can examine the measurement model in Figure 3 more closely. The error terms at the bottom of the figure show the variance unaccounted for in each observable. This can be converted to a proportion by simply calculating the ratio of unexplained variance to total observed variance for the variable. For "Satisfied with Living in 1969" this is .271/.471 = .575, and for the "Get Along on Income" variable the corresponding figure is .296/.905 = .327. Thus 42.5% of the variance (1 - unexplained/total) in the first variable and 67.3% of the variance in the second is explained by the model.[7] These figures are interpretable as reliability coefficients for the items. (See Alwin and Jackson, 1980, for a complete discussion of the relationships between factor analytic and classical true score models.)

The structural model. In this simple model, the structural equations represent the stability of the construct over time. The unstandardized coefficients are approximately .86 for the equation linking T2 to T1 and in the T3-T2 equation. Explained variance in both cases is quite high. These coefficients indicate that the relative order of the respondents on the composite of the two observables is stable. This does not mean that the average value of all respondents has remained the same; there may well have been a constant shift in the mean. LISREL can be used to analyze changes in means over time by including an intercept (Jöreskog, 1979), but for ease of presentation we will assume that all variables have a mean of zero, i.e., that they have been deviated from their means.

## Obtaining Estimates of Coefficients

LISREL estimates the various coefficients in the model using the method of maximum likelihood estimation (MLE). MLE methods are fairly complex, and a complete explanation is substantially beyond the scope of this paper. Wonnocott and Wonnocott (1977) give an excellent introduction to the topic, but the following verbal sketch may give some feel for the problem.

A LISREL model is, in effect, a rather complex set of hypotheses about the causal parameters which give rise to an observed set of variances and covariances. Readers familiar with path analytic concepts will be aware of the rules which generate the observed variances and covariances as a function of the causal estimates in a path diagram. For example, in Figure 3, the covariance between SAT69 and GET69 is the product of the two loadings times the variance of the factor. Thus, the observed covariance of SAT69 and GET69 is 1.0 x 1.742 x .201 (the variance of the factor obtained from the solution) = .350 (the observed covariance of SAT69 and GET69).

In LISREL, as in any path analysis, it is possible to estimate the observed variances and covariances from the causal model. In fact, if one thinks about it, it is the underlying causal process which generates the observed variances and covariances, rather than the reverse. If the posited factor model accounted for the data completely, the loadings and error variances (uniquenesses) could be used to account for the observed variances and covariances exactly. To the extent that the model does not account for the data, the variance-covariance matrix implied by the model will depart from the observed matrix. In MLE, one searches for that set of coefficients that is most likely to have generated the observed variance-covariance matrix, given the hypothesized model.

For example, suppose the model in Figure 3 were expanded to include four indicators of income satisfaction at each of the three points in time. If the hypothesis of a single factor at each time point were correct, the coefficients estimated by LISREL should reproduce the observed variances and covariances within reasonable limits. On the other hand, if a two-factor model were appropriate, LISREL would still produce a set of coefficients, but they would produce a bad approximation to the observed variance-covariance matrix. The coefficients chosen would be those which would make the estimated matrix of observed variances and covariances most likely <u>given</u> the model. This procedure is analogous to that followed in the simple chi-square test of independence in a two-by-two table. In that case, one tests for independence by computing expected cell frequencies, given the marginal distributions and the hypothesis of independence. These estimated cell frequencies are, in fact, MLE, in the sense that they are the most likely values corresponding to the hypothesis, given the marginal distributions on the rows and columns on the table. The usual chi-square test in this case compares the expected frequencies to the observed, and leads us to reject the null hypothesis of independence if the expected and observed do not match within the limits of sampling error.

## Test Statistics and Degrees of Freedom

In LISREL we are interested in comparing the observed matrix of variances and covariances to the estimated matrix generated by the model and asking if they "match" in the same sense that we compare observed and expected frequencies in a simple test of independence. The likelihood ratio statistic (see Jöreskog and Sörbom, 1978, p. 14), which is distributed as chi-square, tells us how likely it is that the observed variance-covariance matrix could have been generated by the hypothesized model. In the model, the coefficients are estimates of population parameters, as is the variance-covariance matrix it implies. We wish to know if the sample (observed) variance-covariance matrix is a good fit to the model.

The chi-square test can be evaluated after the degrees of freedom are computed. To get the degrees of freedom for a particular model, one first counts the number of variances and covariances and then subtracts the number of estimated parameters. For any variance-covariance matrix, the number of unique entries is simply $p(p + 1)/2$, where p is the number of variables. In the present case, p is 6 and $6(6 + 1)/2 = 21$. We are fitting 14 parameters as follows: 1 error term for each observed variable (6 df), the variance of the exogenous unmeasured variable (1 df), 1 factor loading for each factor (since one of the loadings is fixed to a value of 1 in each case) (3 df), stability coefficients from T1 to T2 and from T2 to T3 (2 df) and two errors in equations, one for the T2 factor and one for T3 (2 df). The df are then 21 - 14, or 7. For the model in Figure 3, the chi-square value is 492 which, on 7 df, has a probability substantially less than .0001. Hence we have to reject the hypothesis that the model could have generated the observed variance-covariance matrix. Note that we want the model to fit the data, that is, we want a low value of chi square and a high probability value. Thus, the chi-square test is not particularly informative, since it indicates that the model does not fit the data, but is silent on how this model might compare to other plausible models.

## Modification of the Model

If the chi-square test says that the model does not fit the data, how might we modify the model? There are several possibilities. We might assume that the T1 measure has a direct effect on T3. We might assume that there are, in fact, two factors at each time point. This would be equivalent to assuming that each factor was represented by one item. In both cases, we would be relaxing assumptions about coefficients previously thought to be zero.

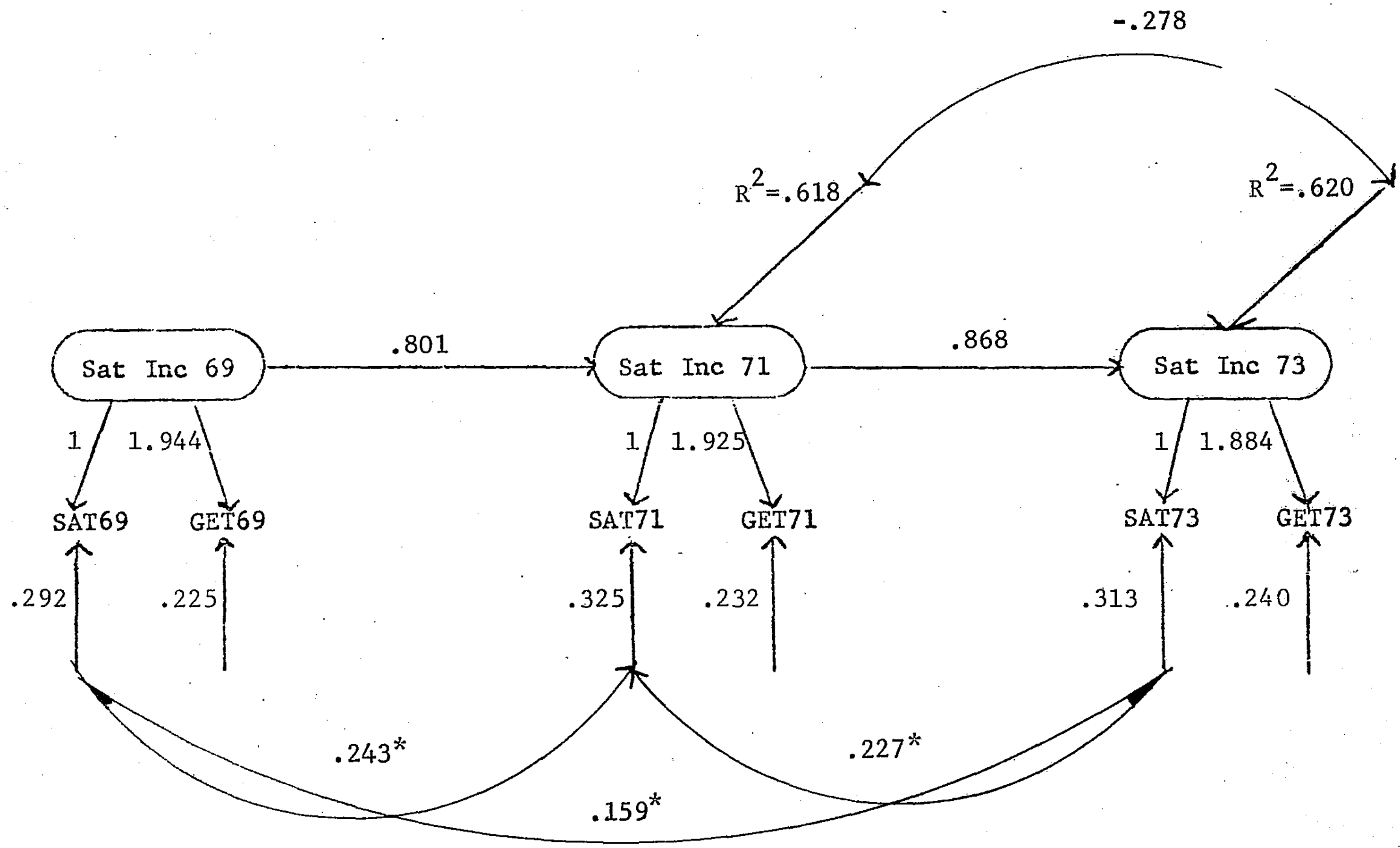
A third class of hypotheses concerns correlated errors. In Figure 3, we explicitly assumed that the errors in variables were unrelated. In survey data, particularly in longitudinal surveys, this may be an unreasonable assumption. It is quite likely that individuals who under- or over-report their score on a particular item will tend to do so again, or that other types of errors, such as misunderstood words or errors caused by the layout of the items, will replicate. These <u>systematic</u> sources of error will cause the correlations among the items to be artificially inflated, and we might want to try to account for correlated errors of measurement with our model.

In addition to errors in variables, Figure 3 also allows for errors in equations, that is, for sources of variance in the predicted outcomes at T2 and T3 which are not due to the independent variables in the model. Again, with panel data, it is not unreasonable to assume that these errors are correlated. If the variance explained is not 100%, because of our omission of a particular explanatory variable, it is likely that the same variable is operating at both points in time. If this is the case, the errors in equations should be correlated.

## A Correlated Error Model

Figure 4 introduces a model in which there are both correlated errors in variables (measurement errors) and correlated errors in equations. Specifically, the model allows the errors on the "Satisifed with Living" item to be correlated over time (but not those on the "Get Along on Income" item) and it allows for the errors in equations to be correlated. We hasten to note that many other assumptions might have been made, and that these are not necessarily the most desirable assumptions to make, substantively speaking. For example, we might have assumed that there is a T1-T2-T3 correlated error (but no T1-T3) for both of the indicators.

FIGURE 4
A THREE-WAVE REPEATED MEASUREMENT MODEL
WITH ITEM-SPECIFIC CORRELATED ERRORS




Note: All coefficients are unstandardized except thosed marked by asterisks, which are correlations. See Appendix A and Table A.1 for definitions and variances of the variables.

$\chi^2$ = .498; df = 3

In any case, this model fits the data extremely well; with 3 df, the chi-square value of .4977 has a probability greater than .9. Note that the four additional parameters estimated are subtracted from the df, i.e., 7 - 4 = 3.

## Comparing Models

We see here that LISREL is most effective when models are compared. What we need to do is fit a series of models under various assumptions and note how the assumptions lead to increments or decrements in the fit between the estimated variance-covariance matrix under the model and the observed data. There is a precise but simple statistical test for the increment or decrement in chi square from one model to another. In the present case, chi square was 492 with 7 df in Figure 3; it declined to .50 with 3 df for the model in Figure 4. If the additional parameters involving correlated errors must all equal to zero, we would assume that the correlated error model would not fit the data any better than a model without such correlations, and that chi square would remain approximately the same. However, if the additional parameters do lead to a better fit, the reduction in chi square should be "significant." Statistically, the difference between two chi square values (likelihood ratio statistics) is itself distributed as chi square with degrees of freedom equal to the difference in df between the two models. For the present case, the reduction of 491.5 in the value of chi-square was accomplished on a reduction in df to 3; the value of 491.5 on 3 df is clearly significant at well beyond the .0001 level. Indeed, the first model did not fit at all and the second fits the data almost perfectly. Such situations are extremely rare in actual practice, of course.

The test just described is only appropriate if one model is a subset of, or "nested" within, another. This implies that, in the second model, parameters which were fixed to some value (usually zero) have been relaxed, but that no parameter which was fixed in the second model has been estimated in the first. If that were the case, the models would not be nested and the chi-square test for increment in fit (which is analogous in many ways to a test of increment of $R^2$ in regression) would not be appropriate. An example of non-nested models appears in the next section.

Introducing Correlated Errors

Figure 4, like Figure 3, reports unstandardized values. However, the numbers on the curved arrows have been converted to correlation coefficients. We find that the errors in equations are correlated at -.278 and that the correlated errors of measurement range from .16 to .24. Although these are small correlations, they are not zero. More important, the estimates of the other coefficients in the model are not independent of these. In statistical jargon, MLE is a full information procedure, that is, each parameter estimate depends on every other estimate.

Introducing this new set of assumptions has fairly serious implications for the remainder of the model. First, the stability coefficients decline, and $R^2$ declines as well. Second, the estimated error variance in GET69 decreases, but the estimated error in SAT69 increases after allowing for correlated errors of measurement. None of these changes is spectacular, but it is necessary to understand that assumptions about measurement error are in fact important for our estimates. Thus, in simple regression analysis, where one is making implicit assumptions to the effect that all variables are measured perfectly and that errors in equations (in the case of simple path analysis) are uncorrelated, the results do in fact depend on those assumptions. LISREL, in many cases, allows one to check the assumptions, and the results are not always comforting.

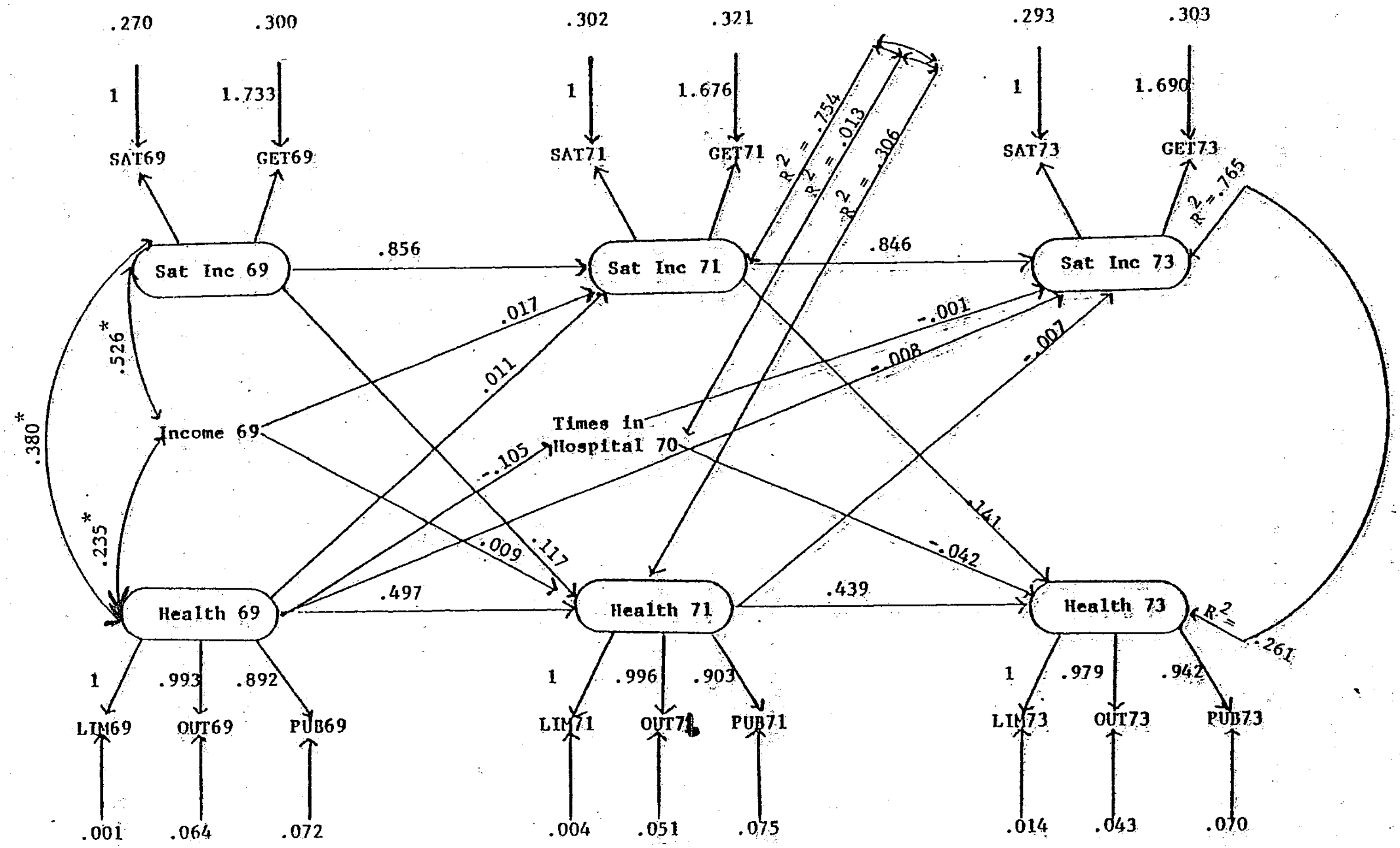
MORE COMPLEX MODELS

Figure 5 introduces a LISREL model of the kind that frequently occurs in the analysis of panel data. Essentially the model contains two variables, health and income satisfaction, at three points in time, along with a baseline measure of family income and an intervening variable indicating number of hospitalizations between the 1969 survey and the one in 1971. The model treats the 1969 variables as exogenous and assumes that the measurement errors in these variables are uncorrelated with measurement errors in the same variables in 1971 and 1973. Errors in equations are allowed to be correlated when no causal linkage is specified among variables measured contemporaneously, e.g., 1971 health and 1971 income satisfaction. Certain causal linkages have been a priori set to zero. For example, we assume no path from "Income" in 1969 to "Times in Hospital" nor is there a path from "Sat Inc 69" to its counterpart in 1973; we assume instead a causal chain via the 1971 measure.

The assumptions made in setting this model up are somewhat unrealistic (certainly we should expect correlated measurement errors), and the high chi-square value of 3138 reflects this lack of empirical

FIGURE 5

A THREE-WAVE MODEL




Note: All coefficients are unstandardized except those marked by an asterisk. See Appendix A and Table A.1 for definitions and variances of the variables.

$\chi^2 = 3137.8$; df = 100.

reality. Still, the results are rather interesting, and we need to discuss them to facilitate comparison with a more realistic model. Again, the figures on the paths are unstandardized coefficients. The coefficients relating the indicators of income satisfaction to the unmeasured variable have the same interpretation as before: they show how much of the variability in the observed variable is due to the unobserved construct, and permit the calculation of reliability coefficients for each item. Similarly, the paths leading from Sat Inc 69 to its 1971 counterpart, and the path from the 1971 measure to the 1973 measure, can be interpreted as stability coefficients. Interestingly, despite the addition of a large number of variables to the model, the stability estimates are about the same: .856 and .846.

At the bottom of the diagram a similar structure appears, relating three indicators of health to an unmeasured construct at parallel points in time. Both the coefficients linking observables to factors and the stability paths can be interpreted as they were in the case of income. Here we find substantially less stability over time, meaning that a person's health status at a given survey is not well predicted by previous status. This would suggest that incidents of ill health occur more or less at random to persons who were previously healthy, although those seriously ill at a particular survey are not likely to be in excellent health at the following data collection point. We also note that the effect of 1969 health status on 1970 hospitalizations is minimal, but that income satisfaction has a fairly large effect on the health measure at the next survey. For example, a one-unit increment in income satisfaction (scaled as SAT69; see Appendix A) leads to a .117 increment in quality of health net of health at time one. That is, those who are satisfied with their income in 1969 tended to be healthier.

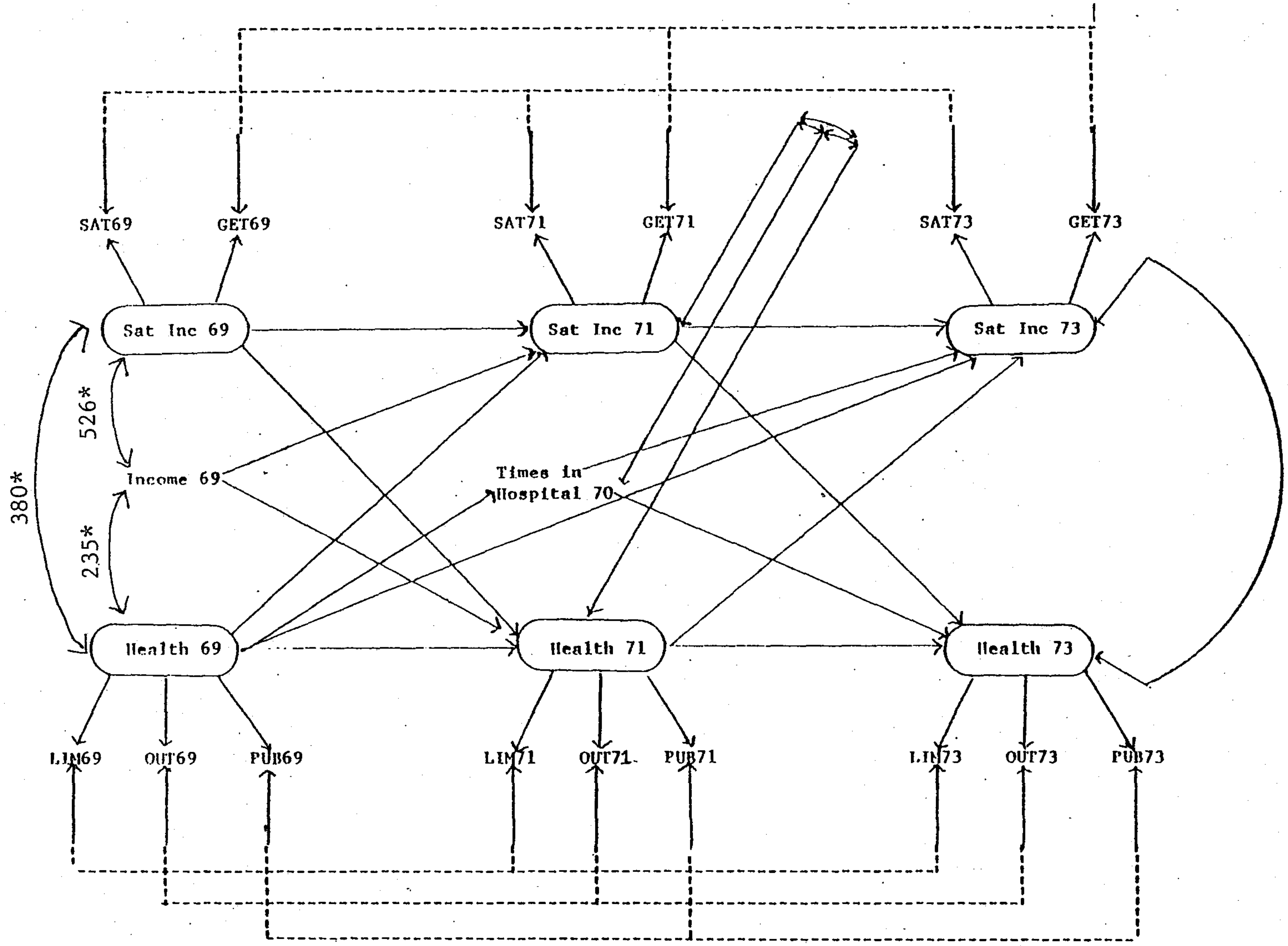
Addition of Correlated Errors

Fi ure 6 shows a model in which item-specific disturbance terms have been allowed to correlate over time in the same manner as Figure 4. For example, errors on LIM69 (LIM indicates effect of health on work; see Appendix A) have been allowed to correlate with errors on LIM71 and LIM73. Similarly, the errors on GET69, GET71, and GET73 have been allowed to correlate mutually. The estimates of this model appear, on the surface, to be quite good; the resulting chi-square value of 1173 indicates a substantial improvement in fit with the expenditure of relatively few degrees of freedom. A reduction of almost exactly 2000 in the value of chi square was obtained by estimating 15 error correlations. However, close examination of the estimated parameters (not shown here) reveals a few anomalies; the estimated error variance of LIM69 is negative, and correlations among error terms approach .9 in some cases.

Aficionados of classical factor analysis will recognize these pathological results as a form of the "Heywood case" (Harman, 1967). The Heywood case occurs when one or more error estimates are less than zero, implying that the factor structure accounts for more than 100% of the variance in the observable. The standard "fix" for this situation is to obtain a factor solution in which the offending error variance is set to zero. This constraint is easily incorporated into a LISREL specification, and a model which restricts the error variance on LIM69 to zero while allowing for error variance on LIM71 and LIM73 as well as correlations between those errors produces a chi-square value of 1613 with 87 df. Unfortunately, it too results in estimated correl tions

FIGURE 6

A THREE-WAVE MODEL WITH CORRELATED ERROR




Note: Coefficients for this model are reported in Table 1. See Appendix A for definitions of variables.

among error terms of .9 or more--clearly pathological results. Thus we see that LISREL models do not automatically yield interpretable results. It is quite possible to get "results" which make no sense.

Given these problems, the task then becomes one of finding a model which fits the data reasonably well while producing plausible results. Since the results for the income satisfaction measures are acceptable, there is no need to change that part of the model specification. The nine items involving the health measures, however, allow for a wide variety of alternate specifications. For example, one could set the error variance of LIM69 to zero and leave the remainder of the model alone. Other alternatives involve equality constraints on error variances and covariances, constraints on error variances only, covariances only, etc. For example, a model which assumes that each error variance is equal over time (that LIM69 = LIM71 = LIM73) and that the item-specific covariances are equal (that cov LIM69, LIM71 = cov LIM69, LIM73 = cov LIM71, LIM73, etc.), thus producing equal error correlations, fits the data reasonably well ($\chi^2$ = 1279), but again, the estimates of the correlations among errors involving LIM are all greater than .9.

The many models which were fitted to the data will not be summarized here; however, we mention two things that became obvious. First, very minor changes in model specification produce very large changes in chi square, and second, the estimates of the causal section of the model change vary little, no matter what is done to the error structure. For example, the stability of the health construct between 1969 and 1971 varied from .462 to .503, depending on the specification of the error structure involving the variances and covariance of LIM, PUB, and OUT, even though the chi-square statistic varied from 1173 to 2037. Put differently, although chi square nearly doubled from one model to another, the parameter estimate changed by less than 10%.

Specification of the final model. The final model has been estimated under the following assumptions concerning errors in variables:

1. For the two indicators of income satisfaction, GET and SAT, cross-time, item-specific errors have been allowed; for example, errors of measurement on GET69 have been allowed to correlate with errors on GET61 and GET73.
2. LIM, the first indicator for the health construct, has been forced to zero error variance at all three time points.
3. Errors in OUT and PUB, the two remaining indicators of the health construct, have been allowed to correlate in the same manner as the indicators of income satisfaction, but in addition we assume that error variances and covariances are equal over time. For example, we assume that the proportion of unexplained variance in PUB is the same in 1969, 1971, and 1973, and that the covariances of PUB are the same for (69,71), (69,73), and (71,73). Note that if we assumed equal covariances but unequal variances it would not be the case that the correlations would be equal over time. Thus the model, as estimated, is a slight variant of that shown in Figure 6.

## Results

The model yields a chi-square value of 2037 with 99 df, one degree fewer than the model shown in Figure 5. It may seem strange that after estimating several correlated errors in variables we have only spent one degree of freedom, but this is because three error variances on LIM were set to zero and only 2 rather than 6 df were required to estimate the error variances of OUT and PUB, given the equality constraints. Thus, although 8 additional parameters involving error covariances were estimated, the net loss of df is one.

No e that these models are not nested; some parameters which were free in the first model (e.g., the variances of LIM) have now been constrained, while other parameters which were fixed in the first model (e.g., the covariances of SAT) have now been relaxed. Thus, the chi-square test of model differences cannot be used to evaluate the decrement in chi square associated with the move from the model of Figure 5 to that of Figure 6.

Results for the structural equation model. In order to leave Figure 6 as legible as possible, results for this analysis are presented in Table 1. The top panel shows the coefficients for the structural equation part of the model in both metric and standardized form. The middle panel shows the correlated errors in equations with $R^2$ statistics on the diagonal. Finally, the bottom panel shows the measurement model with correlations among the errors of measurement shown as small sub-matrices on the right.[8]

Because this paper is primarily didactic in nature rather than substantive, we will not interpret these results at length, but several features of the analysis are worth noting. First, the income satisfaction construct is clearly more stable than the health measure, paralleling the results in Figure 5. The standardized stability parameter for income satisfaction is substantially larger than the corresponding value for health--.754 vs. .495 from Time 1 to Time 2 with similar results for Time 2 to Time 3.

Table 1

Results Pertaining to Figure 6

A. Structural Equation Model

Dependent Variables

| Independent Variables | $\eta_4$ SAT INC 1971 M[a] | S[a] | $\eta_5$ HOSP 1970 M | S | $\eta_6$ HEALTH 1971 M | S | $\eta_7$ SAT INC 1973 M | S | $\eta_8$ HEALTH 1973 M | S |
|---|---|---|---|---|---|---|---|---|---|---|
| Sat Inc 69 $\eta_1$ | .756 | .754 | | | .123 | .116 | | | | |
| Income 69 $\eta_2$ | .038 | .092 | | | .008 | .018 | | | | |
| Health 69 $\eta_3$ | .039 | .041 | -.107 | -.104 | .503 | .495 | -.005 | -.003 | | |
| Sat Inc 71 $\eta_4$ | | | | | | | .803 | .831 | .176 | .162 |
| Times in Hospital 70 $\eta_5$ | | | | | | | -.002 | -.002 | -.048 | -.047 |
| Health 71 $\eta_6$ | | | | | | | -.003 | .004 | .443 | .432 |

B. Errors in Equations[b]

| | | | | | |
|---|---|---|---|---|---|
| $\eta_4$ | .672 | | | | |
| $\eta_5$ | .097 | .013 | | | |
| $\eta_6$ | .180 | .163 | .309 | | |
| $\eta_7$ | | | | .689 | |
| $\eta_8$ | | | | .187 | .275 |

C. Measurement Model

| | "Factor Loadings"[c] $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $\eta_5$ | $\eta_6$ | $\eta_7$ | $\eta_8$ | Error Structure[d] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SAT69 | 1.00* | | | | | | | | .433 | | |
| SAT71 | | | | 1.00* | | | | | .186 | .406 | |
| SAT73 | | | | | | | 1.00* | | .125 | .169 | .396 |
| GET69 | 1.71 | | | | | | | | .656 | | |
| GET71 | | | | 1.67 | | | | | .198 | .633 | |
| GET73 | | | | | | | 1.67 | | .193 | .148 | .646 |
| LIM69 | | | 1.00* | | | | | | 1.0 | | |
| LIM71 | | | | | | 1.00* | | | 0 | 1.0 | |
| LIM73 | | | | | | | | 1.00 | 0 | 0 | 1.0 |
| OUT69 | | | 1.02 | | | | | | .793 | | |
| OUT71 | | | | | | .69 | | | .194 | .791 | |
| OUT73 | | | | | | | | .93 | .194 | .194 | .755 |
| PUB69 | | | .87 | | | | | | .664 | | |
| PUB71 | | | | | | .89 | | | .126 | .697 | |
| PUB73 | | | | | | | | .88 | .126 | .126 | .712 |
| HOSP70 | | | | | 1.00* | | | | | | |
| INC69 | | 2.00* | | | | | | | | | |

NOTE: $\chi^2 = 2036.65$; df = 99

[a] M = metric, S = standardized ($B$ and $\Gamma$ matrices).

[b] Variables defined above. Entries on diagonal are $R^2$ statistics; off-diagonal entries are correlated errors in equations.

[c] Coefficients for regression of observables on unmeasured variables. Asterisk indicates fixed coefficient ($\Lambda_y$ matrix).

[d] See text, page 33, for explanation of these entries.

Turning to the effects of health on income satisfaction, we find a small but significant effect of 1969 health on 1971 income satisfaction and a much smaller lagged effect from 1969 to 1973. The 1971 health measure effect on 1973 income satisfaction is also negligible. On the other hand, income satisfaction has a reasonably large effect on health, both in the 1969-71 period and in 1971-73. Does the fact that income satisfaction has a positive effect on self-reported health indicate that social and psychological well-being has an impact on health, or have we perhaps misspecified the model in some way? We will not pursue this issue here, except to note that we might try many other specifications of the model involving just these variables, to say nothing of other variables in the survey that reflect events and states earlier in the life cycle.

Looking at the measure of times hospitalized, we find that it is only marginally predicted by the 1969 health measure, that it has essentially no effect on 1973 income satisfaction, and that it is very weakly related to health in 1973. We have only a single indicator of hospitalization, and previous research has shown self-reports in this area to be notoriously unreliable (Cannell and Kahn, 1968). Perhaps the minimal effects are due to measurement error, but even with more reliable measures it is unlikely that we would find large effects.

Panel B of Table 1 shows the error structure of the causal part of the model. $R^2$ for each equation is shown on the diagonal of the matrix and the off-diagonal elements are the correlations of disturbance terms for each equation. The specification of correlated errors in this case is quite simple; we have assumed that variables not causally related (e.g., Sat Inc 71 and Health 71) have correlated errors (interpretable as partial correlations between the two variables, net of exogenous variables), but we have not introduced correlated errors between, say, Health 69 and Health 71. To do so would reduce the estimated stability of the construct, but since parallel restrictions on Sat Inc 69 and Sat Inc 71 would probably reduce its stability proportionally, there is no advantage in doing so. In other words, correlated errors would add little substantive information, since it is the relative stability of the constructs that interests us.

The measurement model. Finally, Panel C of Table 1 shows the "factor structure" of the measures. This matrix has been arranged so that the items indexing a particular construct are in adjacent rows. To the right of the factor matrix, under the heading "error structure," the error correlations and item reliabilities of each block of indicators are shown. Note that even though error variance of LIM, PUB, and OUT are restricted to equality over time, the item reliabilities are not equal, since the observed variance changes over time and reliability is defined as true variance/observed variance.

## SOME CAUTIONARY NOTES

### Statistical Assumptions

The LISREL model is extremely powerful, and as with most things in statistics, one does not get something for nothing. In order to use the model it is necessary to make an important statistical assumption, namely that the data have a multivariate normal distribution--that is, that the joint distribution of the variables is normal. This assumption may be relaxed in some cases, but they are beyond the scope of this paper. The cost of violating the assumption of multivariate normality is not yet known; there is very little information on the robustness of LISREL. For the moment, one must be very careful with skewed distributions.

Identification

In LISREL it is possible to get solutions that are not unique; that is, an entirely different set of parameters (indeed an infinite set) might lead to the same estimates of the variance-covariance matrix of observed variables. Such a model is said to be underidentified. For example, the classical factor analysis model is underidentified because the solution may be rotated to a different set of loadings that implies exactly the same set of interitem correlations. In CFA, one specifies certain coefficients to be zero, and under certain conditions these restrictions are sufficient to identify the model so that it is not rotatable. Any rotation would result in a different set of estimates and a different fit of the model to the data, i.e., a different value of chi square. As Joreskog and Sorbom (1978) note, general rules for identification are extremely difficult to give. There are several carefully worked out examples in the LISREL manual, but one needs to develop a feel for the problem. LISREL checks for identification and will print a message when the model is underidentified, but the fact that a solution is reached is not an absolute guarantee that the model is identified.[9] The identification problem is particularly acute when one is trying to estimate feedback loops and/or correlated errors in equations, and novice users should not attempt these specifications.

## Starting Values

LISREL requires the user to provide a set of starting values, one for each of the parameters to be estimated. In theory, the program should reach a solution from any set of starting values; however, the translation from mathematical theory to actual application requires particular computing algorithms and on occasion the program may "blow up" because of bad starting values. In general, the program is more efficient the closer the initial solution (i.e., the set of start values) is to the actual estimates. Bad start values can lead to very costly and unsuccessful attempts to fit models. In some cases, the program will be unable to proceed to a solution given the initial estimates. On occasion, the initial estimates may lead the program to print a message indicating that the model is underidentified when in fact this is not the case. Choosing good start values requires a clear understanding of the various parameters estimated by the program and an intimate knowledge of one's data. On rare occasions the program may reach a "local" rather than a "global" minimum in its attempt to reach a solution. As a precaution, one should always reestimate final models using somewhat different starting values.

## Sample Size and Statistical Power

Our results for model 6 showed that even with a fairly complex error structure the model did not at all fit the data in the statistical sense

($\chi^2$ = 2037, df = 99). For any reasonably large sample size this will almost always be the case. That is, it is almost impossible to find a model which fits the data well, in the statistical sense, without introducing many more parameters than are appropriate from the standpoint of parsimony and simplicity.[10] Several authors have advanced criteria other than chi square for assessing the fit of the model, e.g., the ratio of chi square to df (Wheaton et al., 1977), or examination of residuals. As Jöreskog has pointed out in numerous publications (e.g., Jöreskog, 1979), it is the relative fit of models with which one is most concerned, that is, changes in chi square from one model to another. The likelihood ratio (chi-square) statistic is extremely powerful and, particularly with large sample sizes, will detect even small departures from an exact fit to the observed data. Thus, the analyst must be appreciative of the role of sample size in assessing the model. This problem is not unique to LISREL, however; it is endemic in the social sciences, as Cohen (1969) has pointed out.

## SUMMARY

The purpose of this paper has been to aid researchers in the analysis of longitudinal data, an issue especially pertinent to those engaged in the study of aging. The LISREL model has been presented as a particularly powerful analytical tool. To summarize, it has allowed us to theoretically order our model to incorporate intervening variables, to fix certain paths at zero, and to estimate the effect of this on the model. An important point is that we are not forced to make the assumptions that the variables are measured perfectly, or that the errors in variables are not correlated over time.

Substantively, the relative assessment of income shows a fair amount of stability across time. However, this is not true of good health, and leaves open to further investigation the question of the predictors of good health in old age. One of these factors appears to be concern over income, as expressed by the relationship of relative income to health at each time period. This appears to increase with the passage of time.

The technique presented in the paper should help researchers further to test hypotheses about the effects of variables on one another through time, thus disentangling to some extent the <u>process</u> of aging.

# NOTES

[1]In addition to the references cited in the text, good introductions to path analysis can be found in Duncan (1975), Kerlinger and Pedhazur (1973), and Asher (1976). There are numerous introductions to factor analysis. One of the best is Rummel (1967). For an application of simple path analysis to a problem in gerontology, see Henretta and Campbell (1976).

[2]The intervening variable could, of course, be treated as a time one variable; that is, it could be simply correlated with T1, rather than causally related to it.

[3]Correlated measurement errors can be seen as a special case of an omitted variable.

[4]A factor may be defined by one and only one indicator; that is, the "unmeasured variable" is taken to be equal to the observable indicator.

[5]Many readers trained in classical factor analysis will be uncomfortable with the assumption that a factor can be defined by one or two variables, adhering instead to the "rule" that a factor must be defined by at least three variables. There is no such "requirement" in CFA; the only question is whether the model fits the data, an assessment which can be made using techniques discussed below.

[6]These should not be confused with factor score coefficients, which allow prediction of the factors from observables.

[7]It may appear that something has been obtained here for nothing. There are only two variables on the factor and one observable covariance from which we have estimated two factor loadings and the corresponding unique variance. This is possible only because we actually have six

variables and three factors, from which we are estimating 14 parameters on the basis of 21 entries in the observed variance-covariance matrix. A more complete explanation appears below.

[8]These standardized values are not provided by the program and are obtained by hand from the variances-covariance matrix of errors in variables. Before reporting LISREL's standardized results, be sure to check the formulae given for standardization in the LISREL manual because in some cases they differ from conventional standardizations.

[9]For mathematical reasons too complex to go into here, the identification check is almost certainly accurate. Although the program will catch all underidentified models, it will, on occasion, report that a model is underidentified when in fact it is not, due to starting values (see next section of text, "Starting Values").

[10]This phenomenon is by no means unique to the LISREL model. Social scientists rarely acknowledge the role of statistical power in their decision-making. See Cohen (1969) for an excellent discussion of this point.

# APPENDIX A

## Observable Variables and Unmeasured Constructs

A. Satisfaction with Income is an unmeasured construct with two indicators:

1. SAT69, SAT71, SAT73

   Are you satisfied with the way you are living?

   4 = More than satisfied
   3 = Satisfied
   2 = Less than satisfied
   1 = Very unsatisfied

2. GET69, GET71, GET73

   Ability to get along on income

   4 = Always have money left over
   3 = Have enough with a little left over sometimes
   2 = Have just enough, no more
   1 = Can't make ends meet

B. Health is an unmeasured construct with three indicators:

1. LIM69, LIM71, LIM73

   Does health limit the kind of work you do?

   2 = No
   1 = Yes

2. OUT69, OUT71, OUT73

   Are you able to leave the house without help?

   3 = No limitation
   2 = Yes, though health limits work, can leave house without help
   1 = No

3. PUB69, PUB71, PUB73

   Are you able to use public transportation without help?

   3 = No limitation
   2 = Yes, though health limits work, can use public transportation without help
   1 = No

C. Number of times in hospital (HOS70) is measured with a single indicator.

D. 1969 household income (INC69) is a single indicator of log income from all sources

Table A.1. Variables, Means, Standard Deviations, Variances, and Correlations

| | SAT69 | GET69 | INC69 | LIM69 | OUT69 | PUB69 | SAT71 | GET71 | HOS70 | LIM71 | OUT71 | PUB71 | SAT73 | GET73 | LIM73 | OUT73 | PUB73 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SAT69 | 1.000 | | | | | | | | | | | | | | | | |
| GET69 | 0.535 | 1.000 | | | | | | | | | | | | | | | |
| INC69 | 0.327 | 0.436 | 1.000 | | | | | | | | | | | | | | |
| LIM69 | 0.282 | 0.298 | 0.236 | 1.000 | | | | | | | | | | | | | |
| OUT69 | 0.213 | 0.229 | 0.201 | 0.882 | 1.000 | | | | | | | | | | | | |
| PUB69 | 0.221 | 0.237 | 0.211 | 0.833 | 0.741 | 1.000 | | | | | | | | | | | |
| SAT71 | 0.445 | 0.411 | 0.290 | 0.229 | 0.173 | 0.185 | 1.000 | | | | | | | | | | |
| GET71 | 0.419 | 0.590 | 0.401 | 0.262 | 0.208 | 0.211 | 0.522 | 1.000 | | | | | | | | | |
| HOS70 | −0.039 | −0.041 | −0.020 | −0.102 | −0.074 | −0.079 | −0.077 | −0.070 | 1.000 | | | | | | | | |
| LIM71 | 0.238 | 0.244 | 0.195 | 0.543 | 0.462 | 0.449 | 0.263 | 0.276 | −0.191 | 1.000 | | | | | | | |
| OUT71 | 0.160 | 0.171 | 0.158 | 0.419 | 0.393 | 0.367 | 0.196 | 0.216 | −0.158 | 0.888 | 1.000 | | | | | | |
| PUB71 | 0.175 | 0.177 | 0.159 | 0.413 | 0.375 | 0.379 | 0.198 | 0.217 | −0.162 | 0.833 | 0.779 | 1.000 | | | | | |
| SAT73 | 0.361 | 0.363 | 0.233 | 0.196 | 0.166 | 0.158 | 0.428 | 0.412 | −0.049 | 0.216 | 0.170 | 0.172 | 1.000 | | | | |
| GET73 | 0.374 | 0.521 | 0.356 | 0.245 | 0.198 | 0.207 | 0.408 | 0.582 | −0.069 | 0.233 | 0.185 | 0.186 | 0.507 | 1.000 | | | |
| LIM73 | 0.199 | 0.237 | 0.168 | 0.459 | 0.392 | 0.390 | 0.214 | 0.236 | −0.144 | 0.502 | 0.428 | 0.418 | 0.246 | 0.277 | 1.000 | | |
| OUT73 | 0.122 | 0.146 | 0.126 | 0.313 | 0.292 | 0.279 | 0.141 | 0.159 | −0.111 | 0.385 | 0.382 | 0.370 | 0.195 | 0.215 | 0.886 | 1.000 | |
| PUB73 | 0.128 | 0.153 | 0.128 | 0.308 | 0.279 | 0.284 | 0.143 | 0.160 | −0.104 | 0.376 | 0.366 | 0.371 | 0.189 | 0.214 | 0.833 | 0.807 | 1.000 |
| Mean | 2.781 | 2.451 | 8.531 | 1.661 | 2.663 | 2.710 | 2.803 | 2.508 | 0.156 | 1.636 | 2.679 | 2.699 | 2.780 | 2.465 | 1.578 | 2.660 | 2.672 |
| SD | 0.687 | 0.951 | 1.086 | 0.473 | 0.534 | 0.485 | 0.715 | 0.961 | 0.487 | 0.481 | 0.513 | 0.511 | 0.698 | 0.925 | 0.494 | 0.512 | 0.523 |
| Variance | 0.472 | 0.904 | 1.179 | 0.224 | 0.285 | 0.235 | 0.511 | 0.924 | 0.237 | 0.231 | 0.263 | 0.261 | 0.487 | 0.856 | 0.244 | 0.262 | 0.273 |

# APPENDIX B

## A Primer on LISREL Notation

Many otherwise well-motivated readers of the LISREL literature are put off by its complex notation. All told, there are eight separate matrices to keep track of, each with its own Greek letter, and several other Greek letters are used for other purposes. The purpose of this appendix is to try to organize this notation so as to make it easier to learn.

1. The notation distinguishes between exogenous and endogenous variables on the one hand, and between unobserved constructs and observable variables on the other. The tabulation below shows the notation.

| | Observed Variables | Unobserved Variables |
|---|---|---|
| Exogenous | X | $\xi$ (χi) |
| Endogenous | Y | $\eta$ (eta) |

In Figure 5, the unmeasured income satisfaction measure at time 1 (Sat Inc 69) would thus normally be designated by $\xi$ (χi), as would Health 69. The measured indicators of these variables would be designated with the letter X to indicate that these variables are exogenous--that is, no paths lead to χi from other variables in the model. Income 69 has a single indicator. Nonetheless, we distinguish between the unmeasured construct ($\xi$) and its indicator (X). Sat Inc 73, being endogenous, would be designated by $\eta$ (eta) and its indicators by Y.

2. The notation distinguishes between the causal or structural equation part of the model and the measurement part. The causal part describes the relationships among the unmeasured variables; the measurement part concerns the relationship of the unobserved variables to the observables.

A. The Measurement Model

$\Lambda_y$ (lambda Y) and $\Lambda_x$ (lambda X) contain the coefficients linking the observable y's and x's to the unobservable $\eta$ (eta) and $\xi$ ($\chi$i) respectively. These are regression coefficients in the sense that they give a predicted value of Y or X from the unobserved factor. Each column of the matrix is for an unmeasured variable, each row for a measured variable.

$\Theta_\varepsilon$ (theta epsilon) and $\Theta_\delta$ (theta delta) contain the errors of measurement for predicting the observables from the factors for endogenous and exogenous observables respectively. Each matrix has as many rows and columns as there are observables.

B. The Structural Equation Model

$\Gamma$ (gamma) contains the effects of exogenous unmeasured ($\xi$) ($\chi$i) variables on the endogenous ($\eta$) (eta) variables. Each row of the matrix corresponds to an equation and each column to an exogenous variable.

$\beta$ (beta) contains the effects of endogenous unmeasured ($\eta$) (eta) variables on other unmeasured endogenous variables. There is one row for each equation and one column for each unmeasured endogenous variable.

$\Phi$ (phi) contains the covariances of the unmeasured exogenous variables. It has rows and columns equal to the number of such variables.

$\Psi$ (psi) contains variances and covariances of errors in equations or the amount of variance unexplained by the causal model. There is one row and column in this matrix for each equation.

Note: With one exception, the individual entries in a particular matrix are designated by their lowercase counterpart; e.g., an entry in $\Gamma$ is $\gamma$. The exception is $\Psi$, where specific entries are designated $\zeta$ (zeta).

3. The "all Y" model

LISREL allows correlated errors at measurement among X and Y (exogenous and endogenous) variables, but not between them; e.g., it requires that $\Theta_\varepsilon$ and $\Theta_\delta$ are independent. One can avoid this restriction by defining an "all Y, no X" model, which was done here for figures 5 and 6, thus permitting correlated errors at measurement between exogenous and endogenous constructs.

REFERENCES

Alwin, D. F. and D. J. Jackson. 1980. Measuring error models for survey data: Issues and applications, In K. P. Schuessler (ed.). Sociological methodology, 1980. San Francisco: Jossey-Bass.

Asher, H. B. 1976. Causal modelling. Beverly Hills, Calif.: Sage.

Baltes, P. B. and J. R. Nesselroade. 1979. History and rationale of longitudinal research. In J. R. Nesselroade and P. B. Baltes (eds.), Longitudinal research in the study of behavior and development. New York: Academic Press.

Bock, R. D. 1979. Univariate and multivariate analysis of variance of time structured data. In J. R. Nesselroade and P. B. Baltes (eds.), Longitudinal research in the study of behavior and development. New York: Academic Press.

Cannell, C. F. and R. L. Kahn. 1968. Interviewing. In G. Lindzey and E. Arsonson (eds.), Handbook of social psychology, Vol. 2. Boston: Addison-Wesley.

Cohen, J. 1969. Statistical power analysis for the behavioral sciences. New York: Academic Press.

Cronbach, L. J. and L. Furby. 1970. How we should measure change, or should we? Psychological Bulletin, 74, 68-80.

Duncan, O. D. 1975. Introduction to structural equation models. New York: Academic Press.

Hanushek, E. A. and J. E. Jackson. 1977. Statistical methods for social scientists. New York: Academic Press.

Harman, H. H. 1967. Modern factor analysis. 2nd ed. Chicago: University of Chicago Press.

Heise, D. R. 1969. Separating reliability and stability in test-retest correlation. American Sociological Review, 34, 93-101.

Henretta, J. C. and R. T. Campbell. 1976. Status attainment and status maintenance: A study of social stratification in retirement. American Sociological Review, 41, 981-992.

Irelan, Lola M. 1972. Retirement history study: Introduction. Social Security Bulletin, 35(11), 3-9.

Jöreskog, K. G. 1969. A general approach to confirmatory factor maximum likelihood analysis. Psychometrika, 34, 165-178.

______, 1970. A general method for analysis of covariance structures. Biometrika, 57, 239-243.

______, 1979. Statistical estimation of structural models in longitudinal developmental studies. In J. R. Nesselroade and P. B. Baltes (eds.); Longitudinal research in the study of behavior and development. New York: Academic Press.

Jöreskog, K. G. and D. Sorbom. 1978. LISREL user's guide. Version IV. Chicago: International Educational Services.

Kenny, D. A. 1979. Correlation and causality. New York: Wiley.

Kerlinger, F. N. and E. J. Pedhazut. 1973. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston.

Morgan, J. N. et al. 1972. A panel study of income dynamics: Study design, procedures, available data. Vol. 1. Ann Arbor: Institute for Social Research, University of Michigan.

Nie, N. et al. 1975. SPSS: Statistical package for the social sciences, 2nd ed. New York: McGraw-Hill.

Parnes, H. S. 1975. The national longitudinal surveys: New vistas for labor market research. American Economic Review, 65, 224-249.

Rummel, R. J. 1967. Understanding factor analysis. Journal of Conflict Resolution, 11, 444-480.

Schaie, K. W. and C. Herzog. In press. Longitudinal analysis. In J. Kagan (ed.), Handbook of child development. New York: Norton.

Wheaton, B. et al. 1977. Assessing reliability and stability in panel models. In D. R. Heise (ed.), Sociological methodology, 1977. San Francisco: Jossey-Bass.

Wiley, D. E. and J. A. Wiley. 1970. The estimation of measurement error in panel data. American Sociological Review, 35, 112-117.

Wonnocott, T. H. and R. J. Wonnocott. 1977. Introductory statistics. 3rd ed. New York: Wiley.